Validated By CRFValidator v 1.0.3

Application No: 10724273 Version No: 2.0

Input Set:

Output Set:

Started: 2007-10-15 19:52:39.811
Finished: 2007-10-15 19:52:40.853
Elapsed: 0 hr(s) 0 min(s) 1 sec(s) 42 ms
Total Warnings: 9
Total Errors: 0
No. of SeqIDs Defined: 20
Actual SeqID Count: 20

Error code		Error Description
W	213	Artificial or Unknown found in <213> in SEQ ID (1)
W	213	Artificial or Unknown found in <213> in SEQ ID (2)
W	213	Artificial or Unknown found in <213> in SEQ ID (3)
W	213	Artificial or Unknown found in <213> in SEQ ID (4)
W	213	Artificial or Unknown found in <213> in SEQ ID (5)
W	213	Artificial or Unknown found in <213> in SEQ ID (6)
W	213	Artificial or Unknown found in <213> in SEQ ID (7)
W	213	Artificial or Unknown found in <213> in SEQ ID (8)
W	213	Artificial or Unknown found in <213> in SEQ ID (9)

SEQUENCE LISTING

<110> Palese, Peter
O'Neill, Robert

<120> IDENTIFICATION AND USE OF ANTIVIRAL COMPOUNDS
THAT INHIBIT INTERACTION OF HOST CELL PROTEINS
AND VIRAL PROTEINS REQUIRED FOR VIRAL REPLICATION

<130> 6923-119

<140> 10724273
<141> 2003-11-24

<150> 08/444,994
<151> 1995-05-19

<150> 08/246,583
<151> 1994-05-20

<160> 20

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> NPI-1 specific primer

<400> 1
gcaaagcagg agaaaccac 19

<210> 2
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> nested NPI-1 primer

<400> 2
gggtccatct gatagatatg agag 24

<210> 3
<211> 48
<212> DNA
<213> Artificial Sequence

<220>
<223> 5' RACE anchor primer

<220>
<221> modified_base
<222> 36, 37, 41, 42, 46, 47
<223> n = i

<400> 3
cuacuacuac uaggccacgc gtcgactact acgggnnggg nngggnng 48

<210> 4
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> GSP-I primer

<400> 4
tcctgatgtt gctgtagacg 20

<210> 5
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> GSP-II primer

<400> 5
gcacgactag tatgatttgc 20

<210> 6
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> amino acid #15-22 NS1I-1 protein

<400> 6
Thr Gly Ala Gly Ala Gly Leu Gly
1 5

<210> 7
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> amino acid #163-167 NS1I-1 protein

<400> 7
Tyr Ser Ala Ala Lys
1 5

<210> 8

<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> polylinker sequence

<400> 8
gactggctgg aattccccat ggcgtcc 27

<210> 9
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> polylinker sequence

<400> 9
Asp Trp Leu Glu Phe Pro Met Ala Ser
1 5

<210> 10
<211> 2940
<212> DNA
<213> Homo sapiens

<220>
<223> NPI-1 cDNA

<220>
<221> CDS
<222> (47)...(1663)

<400> 10
ctaacttcag cggtggcacc gggatcggtt gccttgagcc tgaaat atg acc acc 55
Met Thr Thr
1

cca gga aaa gag aac ttt cgc ctg aaa agt tac aag aac aaa tct ctg 103
Pro Gly Lys Glu Asn Phe Arg Leu Lys Ser Tyr Lys Asn Lys Ser Leu
5 10 15

aat ccc gat gag atg cgc agg agg agg gag gaa gaa gga ctg cag tta 151
Asn Pro Asp Glu Met Arg Arg Arg Arg Glu Glu Glu Gly Leu Gln Leu
20 25 30 35

cga aag cag aaa aga gaa gag cag tta ttc aag cgg aga aat gtt gct 199
Arg Lys Gln Lys Arg Glu Glu Gln Leu Phe Lys Arg Arg Asn Val Ala
40 45 50

aca gca gaa gaa gaa aca gaa gaa gaa gtt atg tca gat gga ggc ttt 247
Thr Ala Glu Glu Glu Thr Glu Glu Glu Val Met Ser Asp Gly Gly Phe
55 60 65

cat gag gct cag att agt aac atg gag atg gca cca ggt ggt gtc atc 295
His Glu Ala Gln Ile Ser Asn Met Glu Met Ala Pro Gly Gly Val Ile

```
         70                  75                  80

act tct gac atg att gag atg ata ttt tcc aaa agc cca gag caa cag   343
Thr Ser Asp Met Ile Glu Met Ile Phe Ser Lys Ser Pro Glu Gln Gln
     85                  90                  95

ctt tca gca aca cag aaa ttc agg aag ctg ctt tca aaa gaa cct aac   391
Leu Ser Ala Thr Gln Lys Phe Arg Lys Leu Leu Ser Lys Glu Pro Asn
100                 105                 110                 115

cct cct att gat gaa gtt atc agc aca cca gga gta gtg gcc agg ttt   439
Pro Pro Ile Asp Glu Val Ile Ser Thr Pro Gly Val Val Ala Arg Phe
                120                 125                 130

gtg gag ttc ctc aaa cga aaa gag aat tgt tca ctg cag ttt gaa tca   487
Val Glu Phe Leu Lys Arg Lys Glu Asn Cys Ser Leu Gln Phe Glu Ser
            135                 140                 145

gct tgg gta ctg aca aat att gct tca gga aat tct ctt cag acc cga   535
Ala Trp Val Leu Thr Asn Ile Ala Ser Gly Asn Ser Leu Gln Thr Arg
        150                 155                 160

att gtg att cag gca aga gct gtg ccc atc ttc ata gag ttg ctc agc   583
Ile Val Ile Gln Ala Arg Ala Val Pro Ile Phe Ile Glu Leu Leu Ser
    165                 170                 175

tca gag ttt gaa gat gtc cag gaa cag gca gtc tgg gct ctt ggc aac   631
Ser Glu Phe Glu Asp Val Gln Glu Gln Ala Val Trp Ala Leu Gly Asn
180                 185                 190                 195

att gct gga gat agt acc atg tgc agg gac tat gtc tta gac tgc aat   679
Ile Ala Gly Asp Ser Thr Met Cys Arg Asp Tyr Val Leu Asp Cys Asn
                200                 205                 210

atc ctt ccc cct ctt ttg cag tta ttt tca aag caa aac cgc ctg acc   727
Ile Leu Pro Pro Leu Leu Gln Leu Phe Ser Lys Gln Asn Arg Leu Thr
            215                 220                 225

atg acc cgg aat gca gta tgg gct ttg tct aat ctc tgt aga ggg aaa   775
Met Thr Arg Asn Ala Val Trp Ala Leu Ser Asn Leu Cys Arg Gly Lys
        230                 235                 240

agt cca cct cca gaa ttt gca aag gtt tct cca tgt ctg aat gtg ctt   823
Ser Pro Pro Pro Glu Phe Ala Lys Val Ser Pro Cys Leu Asn Val Leu
    245                 250                 255

tcc tgg ttg ctg ttt gtc agt gac act gat gta ctg gct gat gcc tgc   871
Ser Trp Leu Leu Phe Val Ser Asp Thr Asp Val Leu Ala Asp Ala Cys
260                 265                 270                 275

tgg gcc ctc tca tat cta tca gat gga ccc aat gat aaa att caa gcg   919
Trp Ala Leu Ser Tyr Leu Ser Asp Gly Pro Asn Asp Lys Ile Gln Ala
                280                 285                 290

gtc atc gat gcg gga gta tgt agg aga ctt gtg gaa ctg ctg atg cat   967
Val Ile Asp Ala Gly Val Cys Arg Arg Leu Val Glu Leu Leu Met His
            295                 300                 305
```

```
aat gat tat aaa gtg gtt tct cct gct ttg cga gct gtg gga aac att   1015
Asn Asp Tyr Lys Val Val Ser Pro Ala Leu Arg Ala Val Gly Asn Ile
        310                 315                 320

gtc aca ggg gat gat att cag aca cag gta att ctg aat tgc tca gct   1063
Val Thr Gly Asp Asp Ile Gln Thr Gln Val Ile Leu Asn Cys Ser Ala
    325                 330                 335

ctg cag agt tta ttg cat ttg ctg agt agc cca aag gaa tct atc aaa   1111
Leu Gln Ser Leu Leu His Leu Leu Ser Ser Pro Lys Glu Ser Ile Lys
340                 345                 350                 355

aag gaa gca tgt tgg acg ata tct aat att aca gct gga aat agg gca   1159
Lys Glu Ala Cys Trp Thr Ile Ser Asn Ile Thr Ala Gly Asn Arg Ala
                360                 365                 370

cag atc cag act gtg ata gat gcc aac att ttc cca gcc ctc att agt   1207
Gln Ile Gln Thr Val Ile Asp Ala Asn Ile Phe Pro Ala Leu Ile Ser
            375                 380                 385

att tta caa act gct gaa ttt cgg aca aga aaa gaa gca gct tgg gcc   1255
Ile Leu Gln Thr Ala Glu Phe Arg Thr Arg Lys Glu Ala Ala Trp Ala
        390                 395                 400

atc aca aat gca act tct gga gga tca gct gaa cag atc aag tac cta   1303
Ile Thr Asn Ala Thr Ser Gly Gly Ser Ala Glu Gln Ile Lys Tyr Leu
    405                 410                 415

gta gaa ctg ggt tgt atc aag ccg ctc tgt gat ctc ctc acg gtc atg   1351
Val Glu Leu Gly Cys Ile Lys Pro Leu Cys Asp Leu Leu Thr Val Met
420                 425                 430                 435

gac tct aag att gta cag gtt gcc cta aat ggc ttg gaa aat atc ctg   1399
Asp Ser Lys Ile Val Gln Val Ala Leu Asn Gly Leu Glu Asn Ile Leu
                440                 445                 450

agg ctt gga gaa cag gaa gcc aaa agg aac ggc act ggc att aac cct   1447
Arg Leu Gly Glu Gln Glu Ala Lys Arg Asn Gly Thr Gly Ile Asn Pro
            455                 460                 465

tac tgt gct ttg att gaa gaa gct tat ggt ctg gat aaa att gag ttc   1495
Tyr Cys Ala Leu Ile Glu Glu Ala Tyr Gly Leu Asp Lys Ile Glu Phe
        470                 475                 480

tta cag agt cat gaa aac cag gag atc tac caa aag gcc ttt gat ctt   1543
Leu Gln Ser His Glu Asn Gln Glu Ile Tyr Gln Lys Ala Phe Asp Leu
    485                 490                 495

att gag cat tac ttc ggg acc gaa gat gaa gac agc agc att gca ccc   1591
Ile Glu His Tyr Phe Gly Thr Glu Asp Glu Asp Ser Ser Ile Ala Pro
500                 505                 510                 515

cag gtt gac ctt aac cag cag cag tac atc ttc caa cag tgt gag gct   1639
Gln Val Asp Leu Asn Gln Gln Gln Tyr Ile Phe Gln Gln Cys Glu Ala
                520                 525                 530
```

```
cct atg gaa ggt ttc cag ctt tga agcaatactc tgctttcacg tacctgtgct  1693
Pro Met Glu Gly Phe Gln Leu  *
            535

cagaccaggc tacccagtcg agtcctcttg tggagcccac agtcctcatg gagctaactt 1753
ctcaaatgtt ttccataata ctgtttgcgc tcatttgctt gccttgcgca cctgctctct 1813
tacacacatc tggaaaacct ccggctctct gtggtgggat acccttctaa taaaagggta 1873
accagaacgg cccactctct tttacggaaa aatccctagg ctttggagat ccgcacttac 1933
attagagtta tgggaatata cacatattaa tgtggctccc tttttcttgt gggggaataa 1993
aagaggactc ctcctcattc cctttaacat gggggaaaaa actgacatta aaagatgaga 2053
ctaaatcttt atcttgaatt ttacacaact acttacgaca agggagatgt ttagacctgt 2113
tggtatactt cagagtactt ttcatgagtt cttccacagt gaacccttgg attacctggt 2173
ggctttttct agccagattg cattaatcct tactgagatt ggatggtttt ctttcctcta 2233
ttggcgccat tcttcagata ttaaagttaa accatccact ccctcacctt cagccttcag 2293
tgaatgtgct ttctagttgt caggaatgct gaagaattaa cactttgact cctaaatgtg 2353
atactggtgg gtaagagcag ggcacattta atttgttcgc ttttgcttct ctttggtctg 2413
ggcacattta atttgttcgc ttttgcttct ctttggtctt ttcgaatact tagtaatcga 2473
aaaccatatc ctgtaattta ataaaaaaaa ctaaggacga aaaaacccct ccaattttcc 2533
caaatgcaat cagtgtaact aggggctgtg tttctgcatt aaaataaatg tttcaggctt 2593
tgtggtcctg atcaaggtcc tcattaaaaa attggagttc accctaggct tttcccctct 2653
gtgactggca gataacacat acttttgaaa gtaactttgg gatttttttt cttaggtgca 2713
gctcgattct aatcttttca tgctgcacac gattccttta atcgatagca tccttatctg 2773
aaagaaataa ccatcttctc aacatgacct gcttaaccca aataagaaca gtgatcttat 2833
aacctcattg tttcctaatc tattttattt catctcctgc tagtactgtg ccgcttcccc 2893
ctccccccac acaaaataaa aacagtatct cgcttctggc tcatttt               2940

<210> 11
<211> 538
<212> PRT
<213> Homo sapiens

<220>
<223> deduced protein sequence from NPI-1 cDNA

<400> 11
Met Thr Thr Pro Gly Lys Glu Asn Phe Arg Leu Lys Ser Tyr Lys Asn
 1               5                  10                  15
Lys Ser Leu Asn Pro Asp Glu Met Arg Arg Arg Arg Glu Glu Glu Gly
            20                  25                  30
Leu Gln Leu Arg Lys Gln Lys Arg Glu Glu Gln Leu Phe Lys Arg Arg
        35                  40                  45
Asn Val Ala Thr Ala Glu Glu Glu Thr Glu Glu Glu Val Met Ser Asp
    50                  55                  60
Gly Gly Phe His Glu Ala Gln Ile Ser Asn Met Glu Met Ala Pro Gly
65                  70                  75                  80
Gly Val Ile Thr Ser Asp Met Ile Glu Met Ile Phe Ser Lys Ser Pro
                85                  90                  95
Glu Gln Gln Leu Ser Ala Thr Gln Lys Phe Arg Lys Leu Leu Ser Lys
            100                 105                 110
Glu Pro Asn Pro Pro Ile Asp Glu Val Ile Ser Thr Pro Gly Val Val
        115                 120                 125
Ala Arg Phe Val Glu Phe Leu Lys Arg Lys Glu Asn Cys Ser Leu Gln
    130                 135                 140
Phe Glu Ser Ala Trp Val Leu Thr Asn Ile Ala Ser Gly Asn Ser Leu
145                 150                 155                 160
Gln Thr Arg Ile Val Ile Gln Ala Arg Ala Val Pro Ile Phe Ile Glu
                165                 170                 175
```

```
Leu Leu Ser Ser Glu Phe Glu Asp Val Gln Glu Gln Ala Val Trp Ala
            180                 185                 190
Leu Gly Asn Ile Ala Gly Asp Ser Thr Met Cys Arg Asp Tyr Val Leu
        195                 200                 205
Asp Cys Asn Ile Leu Pro Pro Leu Leu Gln Leu Phe Ser Lys Gln Asn
    210                 215                 220
Arg Leu Thr Met Thr Arg Asn Ala Val Trp Ala Leu Ser Asn Leu Cys
225                 230                 235                 240
Arg Gly Lys Ser Pro Pro Pro Glu Phe Ala Lys Val Ser Pro Cys Leu
                245                 250                 255
Asn Val Leu Ser Trp Leu Leu Phe Val Ser Asp Thr Asp Val Leu Ala
            260                 265                 270
Asp Ala Cys Trp Ala Leu Ser Tyr Leu Ser Asp Gly Pro Asn Asp Lys
        275                 280                 285
Ile Gln Ala Val Ile Asp Ala Gly Val Cys Arg Arg Leu Val Glu Leu
    290                 295                 300
Leu Met His Asn Asp Tyr Lys Val Val Ser Pro Ala Leu Arg Ala Val
305                 310                 315                 320
Gly Asn Ile Val Thr Gly Asp Asp Ile Gln Thr Gln Val Ile Leu Asn
                325                 330                 335
Cys Ser Ala Leu Gln Ser Leu Leu His Leu Leu Ser Ser Pro Lys Glu
            340                 345                 350
Ser Ile Lys Lys Glu Ala Cys Trp Thr Ile Ser Asn Ile Thr Ala Gly
        355                 360                 365
Asn Arg Ala Gln Ile Gln Thr Val Ile Asp Ala Asn Ile Phe Pro Ala
    370                 375                 380
Leu Ile Ser Ile Leu Gln Thr Ala Glu Phe Arg Thr Arg Lys Glu Ala
385                 390                 395                 400
Ala Trp Ala Ile Thr Asn Ala Thr Ser Gly Gly Ser Ala Glu Gln Ile
                405                 410                 415
Lys Tyr Leu Val Glu Leu Gly Cys Ile Lys Pro Leu Cys Asp Leu Leu
            420                 425                 430
Thr Val Met Asp Ser Lys Ile Val Gln Val Ala Leu Asn Gly Leu Glu
        435                 440                 445
Asn Ile Leu Arg Leu Gly Glu Gln Glu Ala Lys Arg Asn Gly Thr Gly
    450                 455                 460
Ile Asn Pro Tyr Cys Ala Leu Ile Glu Glu Ala Tyr Gly Leu Asp Lys
465                 470                 475                 480
Ile Glu Phe Leu Gln Ser His Glu Asn Gln Glu Ile Tyr Gln Lys Ala
                485                 490                 495
Phe Asp Leu Ile Glu His Tyr Phe Gly Thr Glu Asp Glu Asp Ser Ser
            500                 505                 510
Ile Ala Pro Gln Val Asp Leu Asn Gln Gln Gln Tyr Ile Phe Gln Gln
        515                 520                 525
Cys Glu Ala Pro Met Glu Gly Phe Gln Leu
    530                 535
```

<210> 12
<211> 542
<212> PRT
<213> Homo sapiens

<220>
<223> SRP1 protein

<400> 12

```
Met Asp Asn Gly Thr Asp Ser Ser Thr Ser Lys Phe Val Pro Glu Tyr
 1               5                  10                  15
Arg Arg Thr Asn Phe Lys Asn Lys Gly Arg Phe Ser Ala Asp Glu Leu
            20                  25                  30
Arg Arg Arg Arg Asp Thr Gln Gln Val Glu Leu Arg Lys Ala Lys Arg
        35                  40                  45
Asp Glu Ala Leu Ala Lys Arg Arg Asn Phe Ile Pro Pro Thr Asp Gly
    50                  55                  60
Ala Asp Ser Asp Glu Glu Asp Glu Ser Ser Val Ser Ala Asp Gln Gln
65                  70                  75                  80
Phe Tyr Ser Gln Leu Gln Gln Glu Leu Pro Gln Met Thr Gln Gln Leu
                85                  90                  95
Asn Ser Asp Asp Met Gln Glu Gln Leu Ser Ala Thr Val Lys Phe Arg
            100                 105                 110
Gln Ile Leu Ser Arg Glu His Arg Pro Pro Ile Asp Val Val Ile Gln
        115                 120                 125
Ala Gly Val Val Pro Arg Leu Val Glu Phe Met Arg Glu Asn Gln Pro
    130                 135                 140
Glu Met Leu Gln Leu Glu Ala Ala Trp Ala Leu Thr Asn Ile Ala Ser
145                 150                 155                 160
Gly Thr Ser Ala Gln Thr Lys Val Val Val Asp Ala Asp Ala Val Pro
                165                 170                 175
Leu Phe Ile Gln Leu Leu Tyr Thr Gly Ser Val Glu Val Lys Glu Gln
            180                 185                 190
Ala Ile Trp Ala Leu Gly Asn Val Ala Gly Asp Ser Thr Asp Tyr Arg
        195                 200                 205
Asp Tyr Val Leu Gln Cys Asn Ala Met Glu Pro Ile Leu Gly Leu Phe
    210                 215                 220
Asn Ser Asn Lys Pro Ser Leu Ile Arg Thr Ala Thr Trp Thr Leu Ser
225                 230                 235                 240
Asn Leu Cys Arg Gly Lys Lys Pro Gln Pro Asp Trp Ser Val Val Ser
                245                 250                 255
Gln Ala Leu Pro Thr Leu Ala Lys Leu Ile Tyr Ser Met Asp Thr Glu
            260                 265                 270
Thr Leu Val Asp Ala Cys Trp Ala Ile Ser Tyr Leu Ser Asp Gly Pro
        275                 280                 285
Gln Glu Ala Ile Gln Ala Val Ile Asp Val Arg Ile Pro Lys Arg Leu
    290                 295                 300
Val Glu Leu Leu Ser His Glu Ser Thr Leu Val Gln Thr Pro Ala Leu
305                 310                 315                 320
Arg Ala Val Gly Asn Ile Val Thr G
```